Town meeting recollection

- Just collected the slides that where more significant to me
- Nothing from Kasemann, as he just summarized the TAB report, which we already discussed at length by mail
- No conclusion, just remainders for our coming discussion

Overview

- U.S. Grid projects
  - Overview of current state in infrastructure, applications, and middleware
- Next steps
  - NSF "Cyberinfrastructure" report
  - U.S. MAGIC Committee
  - Planned "LHC" ITR proposal & GRIDS-2
  - Building a Grid middleware community
- **U.S. involvement in EGEE**
  - **Integration of infrastructure**
  - **Collaboration on middleware**

U.S. Involvement in EGEE (1): Integration of Infrastructure

- EGEE (& U.S. equivalents) to serve science communities with international scope
  - Physics, astronomy, environment, bio, …, …
- We must design for international collaboration & coordination from the start
  - Need to learn more about application needs & implications on Grid design and operation
  - But some things clear, e.g., standards; open and productive coordination; applications
- Strong application community interest in establishing these connections

U.S. Involvement in EGEE (2): Collaboration on Middleware

- With OGSA, EGEE, evolving distributed support structures, etc., stars are aligned for true international cooperation
  - Software: transform GT & Condor into international collaboration (think Linux) on a common base for Grid/distributed computing
  - Testing & support: link EGEE staff & systems into international testing framework: procedures, tools, infrastructure, bi-literal agreements
- Not easy: many opportunities for not-invented-here, start-from-scratch perspectives, and small local changes that lead to diversion!

U.S. Involvement in EGEE (3): Specific Activities

- Explicit collaborative efforts aimed at
  - Integrating U.S. and EGEE resources in support of international science
  - Creation & operation of coordinated international testing and support structure
- Direct engagement of U.S. groups in EGEE work packages
  - Joint development in some cases, e.g., monitoring/operations
  - Establishment and operation of the structures above, to ensure common evaluation, testing, and support

Instruments

- Integrated Projects
- Networks of Excellence
- Specific Targeted Projects
- Coordinated actions
- Support actions

- Integrated Infrastructure Initiatives
- Coordinated actions
- Support actions
- More info on: http://www.cordis.lu/ist/fp6/activities.htm

Separate calls for proposals!

K. Baxevanidis EU

Communication Network Development Call

- 45-47 Million Euros available in the first EU call (Dec 17th, 2002)
- Hard to get the whole budget, we will need to share with one, two, more projects and a lot of competition to be expected (DEISA, AGRIDnet,...)
- Focus on support and integration of already established Grid infrastructures
- Build a Grid production layer on top of the EU RN infrastructure
- No funds for H/W, CS research or application development (in a first approximation)

Specific service activities

- Integration of major national and international Grid infrastructures
- Two tier structure:
  - **1st Tier: Major Grid centres (4-6). Must satisfy minimum level of Grid resources and staffing**
  - **2nd Tier: POPs in all other RN Geant supported countries**
  - **EU resources for doubling the 1st tier centres Grid support staff, a central operation centre and a distributed call and support centre**
  - **Interface to Geant follow-on project**
  - **Mostly staff and overhead (computer fabrics and storage provided by the partners)**

Joint research activity

- Focus on hardening and re-engineering of Middleware
- Leverage current EU Grid projects and international Grid technology developers (large and established M/W development community)
- A few WPs (3-4) with critical mass in a single geographical center, dedicated WP managers hired by the project and reporting to the project technical management (possible international and industrial participation)
- Quality assurance group, integration, certification and distribution group with industrial quality
- International senior advisory group for project review, long term technology development and direction

Additional activities

- Application support:
  - high level interface and portals
  - user requirements (a la HEPCAL)
  - Multi-science
- CS focused activity:
  - Long term CS issues for production quality Grids

EGEE proposal timeline

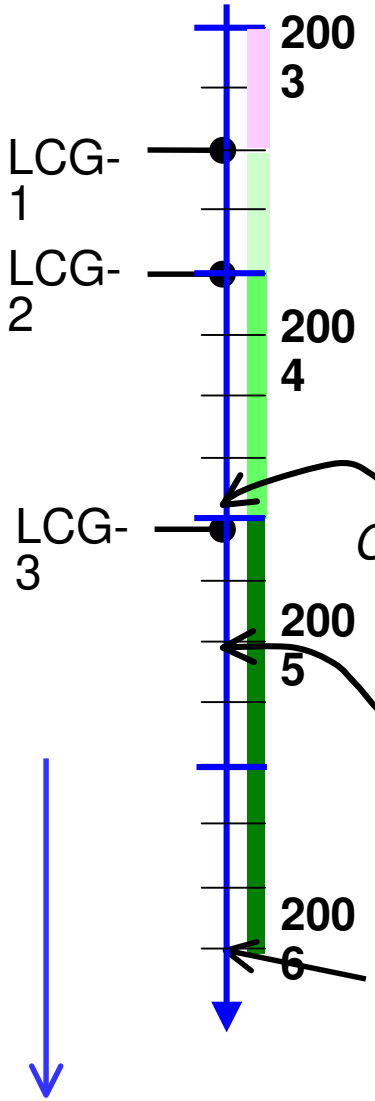
Tentative Schedule

- Draft 1: overall project structure end of February 2003
- Discussed with HEP on a Town Hall meeting on February 22
- Other end-user meetings to be scheduled
- Draft 2: with detailed workpackages end of March 2003
- Final proposal including admin and management end of April 2003
- Submission by May 6th 2003
- First feedback from EU in June-July
- Contract negotiation late summer, fall '03
- Contract signature by the end of '03
- Start of project Q1-Q2 '04

EGEE proposal timeline

- 15/1–5/2 Executive Committee establishes TAB
- 5/2–20/2 TAB prepares recommendations to EC
- 22/2 Town Meeting with HEP
- 22/2–28/2 Executive Committee appoints task forces
- 1/3–30/3 Task forces draft WP contents
- 1/4–4/4 EditorialBoard compiles proposaland submits to EC
- 14/4–18/4 EB prepares new draft
- 23/4 Generalopen meeting
- 24/4–2/5 Finalproposalplus signature and admin
- 6/5 finaldeadline

Timeline for the LCG computing service



VDT, EDG tools building up to basic functionality

Stable 1st generation middleware

Developing management, operations tools

More stable 2nd generation middleware

Very stable full function middleware

Acquisition, installation, commissioning of Phase 2 service (for LHC startup)

used for simulated event productions

principal service for LHC data challenges – batch analysis and simulation

Computing model TDRs

validation of computing models

Phase 2 TDR

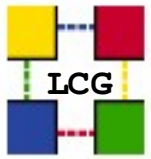
validation of computing service

Phase 2 service in production

Observations

- The middleware tools that we can identify today must be hardened this year and supported into 2005
- It becomes increasingly difficult to introduce new middleware into a stable, production service
- The LHC analysis facility will be fully distributed from day 1 – no site will have more than ~15% of the analysis capacity
- Our minds and mails have been concentrating on the middleware

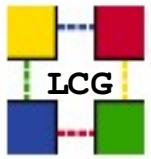
  But there are many (equally difficult) problems to solve in deploying a coherent, productive service for data analysis
- The middleware that will run the service when LHC starts in April 2007, must be deployed at least one year before



Situation Today

- We are still solving basic reliability & functionality problems
  - We still have a long way to go to get to a solid service
  - A solid service in mid-2003 looks ambitious
- We have not yet addressed system level issues
  - How to manage and maintain the Grid as a system providing a high-quality reliable service.
  - Few tools and treatment in current developments of problem determination, error recovery, fault tolerance etc.
- Some of the advanced functionality we will need is only being thought about now
  - Comprehensive data management, SLA's, reservation schemes, interactive use.
- Many many initiatives are underway and more coming

How do we manage the complexity of all this ?



Establishing Priorities

- We need to focus on a model that we can easily explain and understand
- The basic requirements of the 2004 Data Challenges are a good starting point
  - Focus on robust job scheduling, data handling.
- We must make the simple things work well before we expand the scope
- This is not to say that we should not be working on advanced requirements –

  But we must recognise the difference between R&D and providing a service

LCG Grid Middleware Challenges

- Have identified the starting technologies to be deployed
  - Driven pragmatically through the GDB/WG1
  - Initial suppliers – VDT and EDG
- Identify the medium term supply & support strategies
  - Requirements from GAG, advice from STAG
  - Short life-time projects (EDG, Trillium, NMI ...) with unclear continuations
  - We need to see credible supplier projects that focus on product quality, maintainability, support, end-user service
  - What about industrial products
- Work towards future middleware solutions that are coherent, acceptable and supportable
  - Inter-working, standards, ...
    - — essential for evolution
    - — LCG must be mainline, not HEP-special
  - OGSA helps – but will standards emerge on the LCG timescale?
  - or will we need to apply some Super-GLUE?

EGEE as a Solution for LCG

- Middleware
  - EGEE could provide the critical mass
    - to fund **hardened & supported** middleware serving a wider community
    - and form a **global** middleware partnership
  - LCG cannot do this on its own
- Operation
  - funds the operation of a core grid in European countries
  - assuming that –
    - the EGEE infrastructure is integrated with national infrastructures
    - and with US, Asian infrastructures
    - the LCG regional centres are smoothly integrated from the start
    - the aim is a long-term infrastructure
- Simplification
  - LCG (and HEP) can concentrate on physics services and leave middleware and grid operation to someone else
- The long term model is provision of a core Grid service
  - GEANT + NRENs provide a model for this
  - but they had a more tangible deliverable
  - and the model took some time to mature

LCG as a Catalyst for EGEE

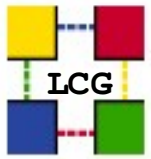
- LCG has a real need
  and a reasonable scale
  and has mature global collaborations of scientists

- LCG must acquire basic middleware for a global science grid

- Solutions for LCG are quite general - readily applicable to other sciences

- LCG will deploy a core grid infrastructure in Europe, America, Asia

  EGEE can build on this
  - to learn how to provide a "production quality" service during LHC preparation (data challenges)
  - exploit the LHC core infrastructure as the foundation for a general science grid



Timing is Critical

- The two-year first phase could deliver just in time (but only just) for LCG
- No time for a slow start-up
- Optimistic approach essential
- Must complete first round – requirements, planning, identify teams, design, ... recruitment .....

  before funding starts to flow

EGEE Priorities (as seen from LCG)

1. Robust industrial-strength middleware, with complete support structure.
   - Functionality at a "basic" level – providing data management, resource scheduling, and information services
   - Part of a Global middleware programme

2. Basic grid operational and support infrastructure – managed as a production service
   - Operations centres, call centres (user support), centres of expertise, system-level support

EGEE & LCG -- opportunities

- EGEE can be an unique opportunity to
  - Build a well operated testbed
  - Provide the necessary personnel to harden (simplify?) the existing MW
  - Provide quality MW using the existing middleware R&D
- The EGEE operated LCG testbed has the potential to provide a focal point for convergence
  - Of different user communities within LHC
  - Of different middleware projects working with LHC
- However such a large project has potential dangers
  - Which are proportionally big!

EGEE & LCG -- caveats

- Divergence
  - US MW projects and US experiment groups MUST “buy in”
    - Cooperate with EGEE / LCG and ensure complementarity
  - Application requirements have to be coordinated and ***controlled***
    - All sciences involved have to get their requirements into the Program of Work
    - But this should not lead to too much divergence
  - All efforts should converge on the same testbed
    - We are seeing with EDG-LCG-1 that the operational interference of several testbeds is destructive on the LHC Regional Centre staff
    - EGEE and LCG-x testbed must coincide (software and hardware) in all sites that belong to both

EGEE & LCG -- caveats

- Rewriting may not be the only option
  - Review current middleware packages with respect to LCG requirements → need repackage, simplify, interoperate, eliminate duplicates ??
  - An architecture would be extremely helpful (components, functionality, API's, protocols)
  - However remember that scrapping software is not failure providing you retain the knowledge -- that would be a good start
- Overhead & timing
  - Planning has to be carefully done as we cannot afford the overhead of running two large projects, supporting two planning/reporting/review processes
  - EGEE timing should be largely in line with LCG timing
- Resources
  - All this has significant costs, EGEE can probably cover it, but only if things are done right from the start
  - If CERN becomes a e-science competence centre this should not be to the detriment of LHC!

EGEE & LCG -- caveats

- Requirements
  - Experiments must make sure that their requirements make it into the Programme of Work
  - But we should be realistic
    - ***Asking for the moon*** will not work, no matter how much manpower is there
  - LHC experiments must be involved in the definition of the workpackages and of their goals
  - But to stand a chance to be heard LHC experiments should speak with a single voice (GAG has been setup for this)

Experiment participation

- Even in the best of all worlds EGEE will draw on experiment resources
  - Installation of software, testing and evaluation of EGEE
  - Necessary participation into the project bodies
  - Collaboration with the different components of the project, in particular MW
- This is not an overhead imposed by EGEE
- It is necessary manpower that we need to build LCG-x
- But EGEE can and must *compensate* for this
  - Failure to secure this manpower would make the participation of experiments into EGEE impossible
  - And therefore would reduce / eliminate the interest of the whole project for LHC

Experiment participation

- We need a "WP8" inside EGEE
- A HEP application work package
  - Some Experiment Independent People (~4) and some additional personnel into the experiments (1-2 people per experiment)
  - Build on the knowledgeable, experienced team within EDG – "loose" cannons
  - Provide support to experiments on the EGEE testbed for installation, evaluation, problem reporting, liaison with the other workpackages
- The EDG experience shows that this is essential
  - Only with such a body the experiments will be able to make the most out of the testbed, properly evaluating it and providing qualified feedback

Relation with GAG

- GAG will continue its work in parallel to EGEE
  - Requirement definition and refinement at a more “abstract” level without getting directly involved with the testbed
  - Look for commonalities in experiment “high level middleware”
  - Official representation in EGEE for LHC requirements
  - Involved in all the phases of the preparation of the EGEE workplan
- It is important that experiments are represented
  - But they must have a common representation
- GAG has been formed to create a common viewpoint of the experiments on GRID

Conclusions

- To avoid dispersion and divergence experiments will have to interact in a highly coherent way with EGEE
  - GAG will act as the LCG forum for developing & monitoring the common requirements → strong i/p to EGEE (& ITR)
- Experiments need extra support to evaluate the testbed and provide qualified feedback
  - EDG has shown that a WP8-like structure is necessary and must be properly manned
  - EDG type “loose cannons” essential for a coherent implementation & evaluation
- EGEE has the potential to be a great success, as we have the expertise and the experience
  - EDG has shown its necessary to have an upfront architecture
  - Essential to have well-described, comprehensive set of use cases